Journal of Finance and Economics
Volume 4, No. 3 (2016), 32-43
ISSN 2291-4951 E-ISSN 2291-496X
Published by Science and Education Centre of North America

Creaming and Dumping: Who on Whom?

Karolina Socha-Dietrich[1]* & Peter Zweifel[2]

[1] Centre of Health Economics Research (COHERE), Department of Business and Economics, University of Southern Denmark

[2] Professor Emeritus, Department of Economics, University of Zurich

*Correspondence: K. Socha-Dietrich, Centre of Health Economics Research (COHERE), Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. Tel. +45 6550 3940; Fax +45 6550 3880; E-mail: k.socha.dietrich@gmail.com

Received: June 18, 2016 Accepted: August 06, 2016 Online Published: September 28, 2016

DOI: 10.12735/jfe.v4n3p32 URL: http://dx.doi.org/10.12735/jfe.v4n3p32

Abstract

In several countries, some healthcare providers combine public service with practice in their own facilities (dual-job practitioners). According to the existing literature, they are viewed as cream-skimming profitable (low-severity) public patients to the benefit of private practice, causing cost of treatment in the public sector to increase. If true, this is particularly problematic when public provider payment is prospective. However, two facts seem to be neglected. First, cream skimming involves effort and thus does not occur in all circumstances. Second, public providers might have an incentive to select patients too, resulting in dumping of the least profitable (high-severity) patients on the private sector. Thus, average cost of treatment in public hospitals does not have to increase, and might even decrease. This paper derives the conditions under which both creaming and dumping are predicted to occur.

JEL Classifications: I11, I18

Keywords: creaming, dumping, waiting lists, public service, dual practice

1. Introduction

Risk selection by healthcare providers may occur in two ways: providers may prefer patients with expected payment greater than expected cost of treatment (cream skimming), and they may get rid of

How to cite this paper: Socha-Dietrich, K., & Zweifel, P. (2016). Creaming and dumping: Who on whom? *Journal of Finance and Economics*, *4*(3), 32-43. http://dx.doi.org/10.12735/jfe.v4n3p32

patients with expected payment less than expected cost (dumping). The incentive to discriminate against high-severity (high-cost) patients arises under prospective payment (e.g. Diagnosis Related Group (DRG) payment), which makes providers bear the risk of excessive cost (Ma, 1994; Newhouse, 1996; Ellis, 1998). This incentive cannot be fully neutralized by risk adjustment because the physician diagnosing a patient is likely to have more information about the patient's future cost than the information contained in the risk-adjustment formula (Newhouse, 1989; Dranove, Kessler, McClellan, & Satterthwaite, 2003). Also, additional effort required in treatment of more difficult and complicated patients might give providers the incentive to avoid them (Newhouse, 1996; Barros & Olivella, 2005). Risk selection both in the guise of cream skimming and dumping is more than a theoretical possibility but has been found in empirical research (Newhouse & Byrne, 1988; Newhouse, 1989; Ellis & McGuire, 1996; Dranove *et al*., 2003).

Risk selection is usually viewed as a problem characterizing market-oriented rather than public healthcare systems (Le Grand, 1991; Ellis, 1998). Even when discrimination against costly patients is illegal in a market-oriented system, some providers may effectively reject them by claiming that they lack the facilities necessary to treat severe cases (Ma, 1994) or by advising medical staff to convince unprofitable patients to seek care elsewhere (Newhouse, 1996). Providers might also cream skim by concentrating on relatively profitable DRGs (Ma, 1994). As a consequence, providers are predicted to differ substantially with regard to the share of expensive patients treated, while severely ill patients face problems of access to health care (Ellis, 1998). Admittedly, extensive regulation in public systems is designed to inhibit exploitation of unpriced risk heterogeneity, causing providers to be tied up with a patient population in a given geographical area, while regulators decide on the range of facilities and services provided.

Yet it has been argued that within public systems, patients might be subject to selection nevertheless. This possibility arises when public providers purchase services from the private sector with the aim of shortening public sector waiting lists in combination with public sector physicians having a second job in a private practice. These dual-job practitioners (DPs) are suspected to manipulate patient transfers in a way that they receive in private practice only the least severe and thus most profitable of public patients (Barros & Olivella, 2005; Gonzalez, 2005). DPs are said to have the incentive because they receive a fixed salary in the public sector, which serves to insulate them from its financial performance. At the same time, they benefit from the low cost of treatment in their private practice which creates a profit margin given prospective payment. This is argued to increase the risk of deficit for public institutions, which suffer from an increase in the average severity and costliness of their patients whereas private institutions appropriate the surplus from treating the low-severity public patients. This reallocation of surplus is seen as inefficient since it undermines the ability of public providers to perform special functions in the public interest, such as education and research.

The present paper re-examines these claims. It proposes a model that takes into account five facts which seem to have been neglected. First, patient selection generally requires effort and time that alternatively could be used for treating patients or for leisure activities. Second, a decrease in the average severity of patients transferred to private facilities by DPs necessarily leads to an increase in the average severity of patients retained in the public sector, causing non-contractible work effort in public facilities to rise. Therefore, single-job practitioners (SPs) who work exclusively in public hospitals may have an interest in dumping high-severity patients on private providers. Third, SPs can be viewed as residual claimants of the surplus generated by public facilities. While this surplus does not accrue to them as a profit, it does provide resources for research, new equipment, and additional facilities (Rickman & McGuire, 1999). Therefore, SPs have a vested interest in keeping the average cost of treatment low in the presence of prospective payment. This interest is shared by hospital managers, who may instruct SPs to dump high-severity patients on private providers as a way to lower average cost of treatment. Fourth, the inclination of DPs to cream skim is predicted in this paper to be weak if their involvement in the private sector is limited or if their total income is high.

Fifth, it is also found that the effectiveness of patient selection must exceed a threshold value for selection to be lucrative.

In sum, the actions of the two groups of providers (DPs and SPs) have opposite effects on the severity of patient cases treated and hence average cost of treatment, resulting in their (partial) cancellation. As a consequence, transfers of public patients to private providers, even in combination with allowing dual job holding, does not necessarily drive up average case severity in the public sector and might in fact decrease it.

The reminder of this paper is structured as follows. Two models of provider behavior are presented in the next section, where cream skimming by DPs versus dumping by SPs is discussed along with optimization by the two types of practitioner. Finally, a conclusion is provided in the last section.

2. Two Models of Provider Behavior

This section contains two models. One depicts the behavior of providers who work exclusively in a public hospital (SPs). The other model applies to providers who combine public service with private practice (DPs). In principle, a third type could be distinguished, i.e. practitioners working exclusively in the private sector who have the right to send patients to a public hospital. Their incentives were analyzed in detail by Zweifel (1981, 1982, & 1985). However, in this paper the hospitalization decision is already made, permitting to focus on SPs and DPs. Both types aim at maximizing their utility, which is increasing in leisure and income. The decision variable is selection effort. For the SP, one has

$$U_p = U_p\left(I_p, L_p\right), \text{ with } \frac{\partial U_p}{\partial I_p} > 0,\ \frac{\partial U_p}{\partial L_p} > 0,\ \frac{\partial^2 U_p}{\partial I_p^2} < 0,\ \frac{\partial^2 U_p}{\partial L_p^2} < 0 \qquad (1)$$

where I_p is income and L_p, leisure. Income I_p in the public sector is assumed to be fixed, e.g. in the guise of a fixed monthly salary. In particular, it does not vary with treatment effort exerted. This assumption may not hold true in all countries; physicians working in public hospitals may in fact be rewarded for their effort. One way is promotion, the other, having private beds. However, latter comes close to dual practice, blurring the distinction between SPs and DPs. Yet, SPs' effort does vary as a function of length of treatment and intensity (stressfulness, respectively) of the work, which in turn depends on the severity of patients treated (Barros & Olivella, 2005). Accordingly, leisure is defined as total time T minus treatment effort S_p (proportional to average severity of cases) minus other types of effort E_p. In particular, E_p can be seen as effort exerted on patient selection,

$$L_p = T - E_p - S_p, E_p > 0 \qquad (2)$$

Turning to the dual practitioner, the utility function of the DP is fully analogous to the one of the SP,

$$U_d = U_d(I_d, L_d), \text{ with } \frac{\partial U_d}{\partial I_d} > 0,\ \frac{\partial U_d}{\partial L_d} > 0,\ \frac{\partial^2 U_d}{\partial I_d^2} < 0,\ \frac{\partial^2 U_d}{\partial L_d^2} < 0 \qquad (3)$$

where I_d and L_d are the DP's income and leisure, respectively. However, his or her income I_d has two components, one arising from involvement in the public sector and the other, from private practice. The first component is given by $k_p I_p$, with k_p being a dimensionless factor, which parameterises the degree of the DP's involvement in the public sector. Together with income from private practice, I_d is given by

$$I_d = k_p I_p + k_d (r - S_d) \tag{4}$$

Here, k_d symbolizes involvement in the private sector. It is worth noticing that, in general, $k_p + k_d \neq 1$. For example, $k_p = 1$ and $k_d > 0$ characterize a DP, who while working full time in the public sector, spends additional time in private practice. The parameter *r* is payments for treating of patients transferred from the public sector. They are reduced by S_d, the patients' mean severity, on the assumption that cost is proportional to severity.

Accordingly, DPs' leisure is given by

$$L_d = T - E_d - (k_p S_p + k_d S_d) > 0 \tag{5}$$

with E_d denoting other effort (again in particular on patient selection) and $(k_p S_p + k_d S_d)$ total treatment effort.[1]

2.1. Cream Skimming by DPs versus Dumping by SPs

It may be worthwhile to recall a few basic facts about the policy of transferring public patients to the private sector. Patients are referred to a public hospital by their general practitioner. In the hospital they initially contact a designated physician, who upon diagnosing them decides whether they should join a waiting list. To be transferred to a private provider, patients need to satisfy certain eligibility criteria such as the expected length of their waiting time. Private facilities contracting with the public payer cannot refuse to treat any of the transferred patients who are part of the agreed contingent (Wiley, 2005; Gonzalez, 2005; Vrangbæk, Østergren, Birk, & Winblad, 2007). Usually, choice between the public and the private facilities is left to eligible patients. Alternatively, patients may leave the choice to the diagnosing physician, who books treatment according to the available operational capacity and the prioritization rule.

The decision variable of the SP is selection effort $E_p \geq 0$. For simplicity, and following Ma (1994), SPs are assumed to be able to predict exactly the amount of resources required by a patient. In this situation, they have an incentive to dump severe cases on private facilities for two reasons. First, as long as effort E_p is smaller than the reduction in treatment effort S_p achieved, they benefit from an increase of leisure [see equation (2)]. Second, SPs can be seen as residual claimants to surplus generated by the public provider which can be used for research, new equipment, or facilities (Rickman & McGuire, 1999). While the existing literature on patient selection does not describe how selection is performed in actual practice, SPs can control the composition of patients they remain in charge of by manipulating transfers to private facilities. One way is to break rules governing the booking of patients for treatment, in particular medical prioritization. Another way is to use their informational advantage to influence patients in their choice between the public and private setting. However, all of this requires effort, denoted by E_p.[2]

Let $e_p > 0$ be the effectiveness of selection effort E_p, decreasing in effort E_p. Therefore, mean

[1] For the sake of simplicity, the relative importance of treatment effort in the private sector and of income derived from it are both reflected by the parameter k_d. A more general formulation does not change results.

[2] The assignment of a patient to the public or private sector is made by a single physician (an SP or a DP). After this point an individual patient is, in general, not reassigned to the other sector, thus barring interaction between the two types of provider.

severity of patients remaining in the public sector decreases by e_pE_p. However, DPs try to reduce the mean severity of patients transferred to their private practice by exerting effort $E_d > 0$, with effectiveness e_d, resulting in a reduction by e_dE_d. Therefore, one has

$$\frac{\partial(e_pE_p)}{\partial E_p}, \frac{\partial(e_dE_d)}{\partial E_d} > 0, \frac{\partial e_p}{\partial E_p}, \frac{\partial e_d}{\partial E_d} < 0 \text{ and } \frac{\partial^2(e_pE_p)}{\partial E_p^2}, \frac{\partial^2(e_dE_d)}{\partial E_d^2} < 0 \tag{6}$$

Moreover, for both the SP and the DP, cream skimming and dumping by the other necessarily leads to an increase in the severity of patients treated. Hence,

$$S_p = S - e_pE_p + e_dE_d \tag{7}$$

$$S_d = S + e_pE_p - e_dE_d \tag{8}$$

where S denotes mean severity in the population.

2.2. Optimization by the Two Types of Practitioner

According to labor economics, dual job holding is a response to income constraints such as limited demand for labor and/or standardized work contracts. If there were no constraints in any of the two jobs, individuals would focus exclusively on the preferred one (Perlman, 1966; Shishko & Rostker, 1976). The public healthcare sector abounds with regulations regarding hours of work, which may explain the prevalence of DPs who combine work in public and private practice. Moreover, in healthcare systems with dominance of public hospital care, demand for private services is limited. In terms of the two models, these constraints justify treating the two involvement parameters k_p and k_d as exogenous, leaving E_p and E_d as the only decision variables of the SP and DP, respectively.

2.2.1. The Single-Job Practitioner

From equation (2), for the SP in the public sector the first-order condition for an interior solution reads,

$$\frac{\mathrm{d}U_p}{\mathrm{d}E_p} = \frac{\partial U_p}{\partial L_p}\frac{\partial L_p}{\partial E_p} = 0, \tag{9}$$

since income I_p does not depend on effort E_p. In view of equations (2) and (7), equation (9) can be rewritten to become

$$\frac{\partial U_p}{\partial L_p}\left(\frac{\partial(e_pE_p)}{\partial E_p} - 1\right) = 0. \tag{10}$$

Since $\partial U_p/\partial L_p > 0$, this boils down to

$$\frac{\partial(e_pE_p)}{\partial E_p} = \frac{\partial e_p}{\partial E_p}E_p + e_p = (\eta_p + 1)e_p = 1 \tag{11}$$

where $\eta_p = \frac{\partial e_p}{\partial E_p}\frac{E_p}{e_p} < 0$ is the elasticity of selection effectiveness w.r.t. selection effort.[3] Therefore,

[3] Also, $\eta_p > -1$, as according to equation (6), $0 > \partial e_p/\partial E_p = e_p(1 + \eta_p)$ and $e_p > 0$.

$e_p > 1$ at an interior optimum, while $e_p \leq 1$ everywhere leads to a boundary optimum $E_p = 0$. As a consequence, the SP has an incentive to select patients only if the efficiency of dumping exceeds the critical value $\hat{e}_p = 1$. Otherwise, the SP makes no effort to dump high-severity patients; only for $e_p > 1$ does the gain from selection outweigh the effort involved.

2.2.2. The Dual-Job Practitioner

In full analogy to equation (9), the first-order condition w.r.t. effort for a DP reads,

$$\frac{\mathrm{d}U_d}{\mathrm{d}E_d} = \frac{\partial U_d}{\partial I_d}\frac{\partial I_d}{\partial E_d} + \frac{\partial U_d}{\partial L_d}\frac{\partial L_d}{\partial E_d} = 0 \tag{12}$$

The distinguishing feature is that for the DP, the mean severity of patients does not only affect leisure, but also income from private practice. In view of equations (3) to (5), and (8), the first-order condition reads,

$$\frac{\partial U_d}{\partial I_d}\left(-k_d\frac{\partial S_d}{\partial E_d}\right) + \frac{\partial U_d}{\partial L_d}\left(-1 - k_p\frac{\partial S_p}{\partial E_p} - k_d\frac{\partial S_d}{\partial E_d}\right) =$$

$$\frac{\partial U_d}{\partial I_d}(-k_d)\left(-\frac{\partial(e_d E_d)}{\partial E_d}\right) + \frac{\partial U_d}{\partial L_d}\left[-1 - k_p\frac{\partial(e_d E_d)}{\partial E_d} - k_d\left(-\frac{\partial(e_d E_d)}{\partial E_d}\right)\right] =$$

$$\frac{\partial U_d}{\partial I_d}k_d\frac{\partial(e_d E_d)}{\partial E_d} + \frac{\partial U_d}{\partial L_d}\left[\left(k_d - k_p\right)\frac{\partial(e_d E_d)}{\partial E_d} - 1\right] = 0. \tag{13}$$

Using the elasticity $\eta_d = \frac{\partial e_d}{\partial E_d}\frac{E_d}{e_d} \leq 0$ but > -1, the marginal rate of substitution between income and leisure $\mu_d = \frac{\partial U_d}{\partial I_d}/\frac{\partial U_d}{\partial L_d} > 0$, and introducing relative involvement in private practice $\rho = k_d/k_p$, one obtains

$$\frac{\partial U_d}{\partial I_d}k_d(\eta_d + 1)e_d + \frac{\partial U_d}{\partial L_d}\left[\left(k_d - k_p\right)(\eta_d + 1)e_d - 1\right] = 0$$

$$\Rightarrow \mu_d k_d(\eta_d + 1)e_d + \left(k_d - k_p\right)(\eta_d + 1)e_d - 1 = 0$$

$$\text{or} \qquad (\eta_d + 1)e_d[(1 + \mu_d)\rho - 1] - \frac{1}{k_p} = 0 \tag{14}$$

It is of particular interest to know when there is positive selection effort by the DP, i.e., when equation (14) has an interior solution. Denoting by E_d^* the optimum value of E_d, one indeed finds that at $\rho = 0$ (no involvement in private practice), $E_d^* = 0$ (an optimum on the boundary), as there is no solution to the first-order condition (14) in view of $\eta_d + 1 > 0$. This is consistent with the fact that a DP without involvement in the private practice is actually an SP. As long as $E_d^* = 0$, e_d and η_d remain unchanged while ρ increases. For sufficiently large ρ, however, the term in square brackets becomes positive and eventually compensates $-\frac{1}{k_p}$ [unless (5) is violated already before]. This means that there exists a ρ above which the DP will exert selection effort.

2.2.3. Does More Involvement in Private Practice Entail an Increase in Selection Effort?

The question to be answered in this section is whether an increase in private practice involvement systematically causes the DP to step up selection effort. In the following an exogenous change $d\rho > 0$ is considered. Admittedly, at least in the long run, ρ is a decision variable. At a given point in time, however, an increase in ρ serves to distinguish between DPs with more and DPs with less relative involvement in private practice. Totally differentiating the first-order condition, one obtains

$$\frac{\partial}{\partial E_d}\frac{\partial U_d}{\partial E_d}\mathrm{d}E_d^* + \frac{\partial}{\partial \rho}\frac{\partial U_d}{\partial E_d}\mathrm{d}r = 0 \qquad (15)$$

The second-order condition for a maximum, $\frac{\partial^2 U_d}{\partial E_d^2} < 0$, is assumed to be satisfied; hence, the sign of $\frac{\mathrm{d}E_d^*}{\mathrm{d}\rho}$ is given by the sign of $\frac{\partial}{\partial \rho}\frac{\partial U_d}{\partial E_d}$ at the optimum. With fixed k_p, one obtains from equation (13)[4]

$$\frac{\partial}{\partial \rho}\frac{1}{k_p}\frac{\partial U_d}{\partial E_d} = \frac{\partial U_d}{\partial I_d}\left(\frac{\rho}{\mu_d}\frac{\partial \mu_d}{\partial \rho} + 1 + \frac{1}{\mu_d}\right)(\eta_d + 1)e_{d.} \qquad (16)$$

All factors on the right-hand side are positive apart from the expression in big parentheses, which has an indefinite sign. The first term in that expression is negative under the reasonable assumption that the DP does not lose money by working in the private sector, $r - S_d > 0$ (an increasing value of ρ with k_p fixed means less leisure, $\frac{1}{k_p}\frac{\partial L_d}{\partial \rho} = -S_d < 0$, and more income, $\frac{1}{k_p}\frac{\partial I_d}{\partial \rho} = r - S_d > 0$). This negative first term can still be compensated by the remaining terms, which are always positive. By assumption $(E_d^* > 0)$, there is a parameter range where equation (14) has a solution. Thus, $\frac{\partial E_d^*}{\partial \rho} \geq 0$ and generally, $\frac{\partial E_d^*}{\partial \rho} > 0$ at least on a finite interval. If there is a solution to the first-order condition, the DP who gets more involved in private practice is predicted to increase selection effort at least over this range of p. Whether there will cease to be solutions of equation (14) for some larger values of ρ, i.e., whether the DP will stop exerting selection effort, depends on the properties of the utility and effectiveness functions. It appears reasonable to assume that μ_d approaches zero as ρ becomes large because, as explained above, this corresponds to more work and more income, i.e., to a lower valuation of income relative to leisure. Then, if equation (14) remains solvable for a large value of ρ, $(\eta_d + 1)e_d = \partial(e_d E_d)/\partial E_d$ must decrease (eventually to zero) to compensate for this increase. Conversely, if the first-order condition is only solvable up to a certain maximum value of ρ, there must be two solutions (they might coincide in a borderline case) for equation (14), with $(\eta_d + 1)e_d$ evaluated at $E_d = 0$.[5]

In sum, the DP is predicted to exert positive selection effort

- only above a minimum value of involvement in the private practice - see discussion of equation (25);
- only if the positivity requirement on leisure is not violated, which can make selection effort entirely unattractive - see equations (5) and (28);
- that increases with increasing involvement in the private practice - equation (16) and

[4] The details of the derivation are provided in appendix A.

[5] Details are provided in appendix B.

(a) starts decreasing and ceases altogether at some maximum involvement,

(b) or keeps increasing if the rate of substitution between income and leisure falls rather slowly with increasing involvement, in which case it is ultimately cut off by (5).

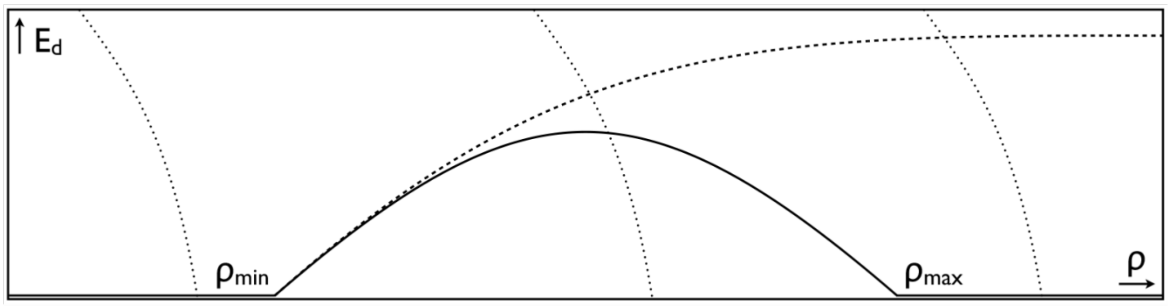

Figure 1. E_d as a function of ρ if equation (14) has solutions for $\rho > \rho_{\min}$

In figure 1, the solid line depicts the case (a), where the solutions stop existing for $\rho > \rho_{\max}$, which happens if μ_d decreases sufficiently fast with increasing ρ. The dashed line represents the case (b), where there is no change of sign in $\partial E_d^*/\partial\rho$ and where there are always solutions to the first-order condition. The dotted lines show three possible cases where $L_d > 0$ is violated. The rightmost dotted line, shows the case where (a) is not affected while in (b) the solution of equation (14) cannot be maintained for (very) large values of ρ. The middle dotted line also affects (a) directly by cutting it of, whereas the leftmost dotted line represents case (b) with no solution that does not violate $L_d > 0$.

The first-order condition can also be interpreted in a different way. For $\left[\frac{\partial U_d}{\partial I_d}k_d + \frac{\partial U_d}{\partial L_d}(k_d - k_p)\right]E_p^*, E_d^* > 0$, there exists a nontrivial solution

$$0 < \left\{\frac{\frac{\partial U_d}{\partial L_d}}{\frac{\partial U_d}{\partial I_d}k_d + \frac{\partial U_d}{\partial L_d}(k_d - k_p)}\right\}_{E_p^*, E_d^*} = \left(\frac{\partial(e_d E_d)}{\partial E_d}\right)_{E_p^*, E_d^*} < e_d | E_d = 0 \tag{17}$$

If $e_d | E_d = 0$ exceeds a finite threshold for cream skimming. The exact functional form of this threshold depends on e_d and its derivative at E_d^*. Due to the positivity requirement for leisure [see equation (28)] - the value of ρ necessary for the DP to become interested in cream skimming may, in fact, become so large that it is above the point where the DP is willing to exert any kind of additional effort. The result is that the DP will not cream skim for any value of ρ.

3. Discussion

The present paper focuses on a public healthcare system where public patients are transferred from public to private providers with the aim of shortening waiting lists in the public sector. The existing literature claims that under such a policy, dual-job practitioners (DPs), who combine public service with private practice, manipulate patient transfers in a way that only the least severe and thus most profitable of public patients are treated in private practice. This is said to result in an increase in average severity and costliness of patients retained in the public sector. The analysis presented in this paper shows this not necessarily to be the case. Patient selection costs time and effort that alternatively could be used for treating patients or for leisure activities. Specifically, there exists a finite threshold for the DP's involvement in private practice below which patient selection is not lucrative. Above that threshold, selection effort is predicted to increase with increasing involvement in private practice under realistic conditions - but only to some degree. A crucial condition is a sufficiently high effectiveness of DPs' cream-skimming effort. Moreover, physicians working exclusively in the public sector (SPs) in turn have an incentive to dump high-severity patients on the private sector. This is because high- severity patients entail disutility due to additional work and an erosion of margins under DRG-based prospective payment.

The analysis suggests that the effects of cream-skimming efforts by DPs and the dumping efforts by SPs on average severity of those patients who remain in the public sector can cancel each other at least to some extent. While it would go beyond the present paper to formally model the net effect of the two types of effort, a preliminary assessment seems possible. In healthcare systems where hospital care is provided predominantly in the public sector, there are many more SPs than DPs. This has a twofold effect for the mean severity and hence treatment cost of patients treated in the public sector. First, DPs have comparatively limited leverage in influencing average severity and cost in the public sector. Second, this means that their cream-skimming effort has low effectiveness, making it less likely that DPs ever engage in patient selection, as shown in this paper.

4. Limitations and Conclusions

This analysis is subject to several limitations. First, neither the DP nor the SP are viewed as having special professional ethics, which might be intrinsic or induced be reputation effects. Second, the two types of provider do not anticipate each other's actions, although they usually can identify each other. A more detailed analysis may therefore call for a game-theoretic formulation. Third, the physician's hospitalization decision is viewed as exogenous; however, it may well depend on his or her SP or DP status. Still, the present contribution calls attention to the fact that cream skimming and dumping are not a one-way street, as is often surmised in the existing literature. A particular case may serve to buttress this statement; it is presented without claim to generality. A private clinic in Germany performed more than 1,000 endoprosthetic knee surgeries in 2013, of which 45 percent are partial knee replacement performed mainly by DPs. This rate is much higher than the national average of 10 percent. While substantially more demanding, this innovative alternative to total knee replacement has the great benefit of leaving the cruciate ligaments intact (Willis-Owen, Brust, Alsop, Miraldo, & Cobb, 2009). At the same time, DRG payment for it is much lower than for conventional total knee replacement. The likely reason for the DPs in the private clinic to attract high-cost patients is reputation: countrywide, primary care physicians prefer to send their patients to it, trusting high-volume providers to perform the surgery with a maximum rate of favorable outcomes.

Acknowledgements

The authors are grateful to the anonymous referee for suggestions and criticism.

References

[1] Barros, P. P., & Olivella, P. (2005). Waiting lists and patient selection. *Journal of Economics and Management Strategy*, *14*(3), 623-646. doi:10.1111/j.1530-9134.2005.00076.x

[2] Dranove, D., Kessler, D., McClellan, M., & Satterthwaite, M. (2003). Is more information better? The effects of "report cards" on health care providers. *Journal of Political Economy*, *111*(3), 555-588. doi:10.1086/374180

[3] Ellis, R. P. (1998). Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *Journal of Health Economics*, *17*(5), 537-555. doi:10.1016/S0167-6296(97)00042-8

[4] Ellis, R. P., & McGuire, T. G. (1996). Hospital response to prospective payment: moral hazard, selection, and practice-style effects. *Journal of Health Economics*, *15*(3), 257-277. doi:10.1016/0167-6296(96)00002-1

[5] Gonzalez, P. (2005). On a policy of transferring public patients to private practice. *Health Economics*, *14*(5), 513-527. doi:10.1002/hec.946

[6] Le Grand, J. (1991). Quasi-markets and social policy. *The Economic Journal*, *101*(408), 1256-1267. doi:10.2307/2234441

[7] Ma, A. C. (1994). Health care payment systems: Cost and quality incentives. *Journal of Economics and Management Strategy*, *3*(1), 93-112. doi:10.1111/j.1430-9134.1994.00093.x

[8] Newhouse, J. P. (1989). Do unprofitable patients face access problems? *Health Care Financing Review*, *11*(2), 32-42.

[9] Newhouse, J. P. (1996). Reimbursing health plans and health providers: Efficiency in production versus selection. *Journal of Economic Literature*, *34*(3), 1236-1263.

[10] Newhouse, J. P., & Byrne, D. J. (1988). Did medicare's prospective payment system cause length of stay to fall? *Journal of Health Economics*, *7*(4), 413-416. doi:10.1016/0167-6296(88)90023-9

[11] Perlman, R. (1966). Observations on overtime and moonlighting. *Southern Economic Journal*, *33*(2), 237-244. doi:10.2307/1055390

[12] Rickman, N., & McGuire, A. (1999). Regulating providers' reimbursement in a mixed market for health care. *Scottish Journal of Political Economy*, *46*(1), 53-71. doi:10.1111/1467-9485.00120

[13] Shishko, R., & Rostker, B. (1976). The economics of multiple job holding. *The American Economic Review*, *66*(3), 298-308.

[14] Vrangbæk, K., Østergren, K., Birk, H. O., & Winblad, U. (2007). Patient reactions to hospital choice in Norway, Denmark, and Sweden. *Health Economics, Policy and Law*, *2*(2), 125-152. doi:10.1017/S174413310700401X

[15] Wiley, M. M. (2005). The Irish health system: Developments in strategy, structure, funding and delivery since 1980. *Health Economics*, *14*(S1), S169-S186. doi:10.1002/hec.1034

[16] Willis-Owen, C. A., Brust, K., Alsop, H., Miraldo, M., & Cobb, J. P. (2009). Unicondylar knee arthroplasty in the UK National Health Service: An analysis of candidacy, outcome and cost efficacy. *The Knee*, *16*(6), 473-478. doi:10.1016/j.knee.2009.04.006

[17] Zweifel, P. (1981). Supplier-induced demand in a model of physician behavior. In J. van der Gaag & M. Perlman (Eds.), *Health, economics, and health economics*. Paper presented at the proceedings of the World Congress on Health Economics, Leiden, The Netherlands, September 1980 (p.245-267). Amsterdam: North Holland.

[18] Zweifel, P. (1982). *Ein ökonomisches modell des arztverhaltens* (Lecture notes in economics and mathematical systems, Vol. 198). Heidelberg: Springer. doi:10.1007/978-3-642-95405-4

[19] Zweifel, P. (1985). Technology in ambulatory medical care: Cost increasing or cost saving? *Social Science and Medicine*, *21*(10), 1139-1151. doi:10.1016/0277-9536(85)90172-8

Appendix A

The first-order condition depends not only on $\rho = k_d/k_p$ but separately on k_d and k_p. This originates from the dependence of the DP's utility on the total amount of effort spent in both sectors combined, to the detriment of leisure. Therefore, in a comparative-static analysis involving ρ, the parameters to be varied must be specified. Here, k_d is this parameter, while k_p is held constant. This corresponds to the substitution $(k_p, k_d) \rightarrow (k_p, \rho)$. In that case, $\partial k_p/\partial\rho = 0$.[6]

Starting from equation (13),

$$\frac{\mathrm{d}U_d}{\mathrm{d}E_d} = \frac{\partial U_d}{\partial I_d} k_d \frac{\partial(e_d E_d)}{\partial E_d} + \frac{\partial U_d}{\partial L_d}\left[\left(k_d - k_p\right)\frac{\partial(e_d E_d)}{\partial E_d} - 1\right] \tag{18}$$

divide by k_p to obtain

$$\frac{1}{k_p}\frac{\mathrm{d}U_d}{\mathrm{d}E_d} = \frac{\partial U_d}{\partial I_d}\frac{k}{k_p}\frac{\partial(e_d E_d)}{\partial E_d} + \frac{\partial U_d}{\partial L_d}\left[\left(\frac{k_d}{k_p} - 1\right)\frac{\partial(e_d E_d)}{\partial E_d} - \frac{1}{k_p}\right]. \tag{19}$$

Next, k_d is replaced by $k_p\rho$ in accordance with the above substitution rule,

$$\frac{1}{k_p}\frac{\mathrm{d}U_d}{\mathrm{d}E_d} = \frac{\partial U_d}{\partial I_d}\rho\frac{\partial(e_d E_d)}{\partial E_d} + \frac{\partial U_d}{\partial L_d}\left[(\rho - 1)\frac{\partial(e_d E_d)}{\partial E_d} - \frac{1}{k_p}\right]. \tag{20}$$

Using $\mu_d = \frac{\partial U_d}{\partial I_d}/\frac{\partial U_d}{\partial L_d}$ yields

$$\frac{1}{k_p}\frac{\mathrm{d}U_d}{\mathrm{d}E_d} = \frac{\partial U_d}{\partial I_d}\left\{\rho\frac{\partial(e_d E_d)}{\partial E_d} + \frac{1}{\mu_d}\left[(\rho - 1)\frac{\partial(e_d E_d)}{\partial E_d} - \frac{1}{k_p}\right]\right\} \tag{21}$$

and taking the derivative w.r.t. ρ,

$$\frac{\partial}{\partial\rho}\frac{1}{k_p}\frac{\mathrm{d}U_d}{\mathrm{d}E_d} = \frac{\partial U_d}{\partial I_d}\left\{\frac{\partial(e_d E_d)}{\partial E_d} + \frac{1}{\mu_d}\frac{\partial(e_d E_d)}{\partial E_d} + \left(\frac{\partial}{\partial\rho}\frac{\partial U_d}{\partial I_d}\right)\underbrace{\left[\frac{\frac{\mathrm{d}U_d}{\mathrm{d}E_d}}{k_p(\frac{\partial U_d}{\partial I_d})^2}\right]}_{\overset{\text{f.o.c.}}{=}0} - \right.$$
$$\left. - \frac{1}{\mu_d^2}\frac{\partial\mu_d}{\partial\rho}\underbrace{\left[(\rho - 1)\frac{\partial(e_d E_d)}{\partial E_d} - \frac{1}{k_p}\right]}_{\overset{\text{f.o.c.}}{=}-\mu_d\rho\frac{\partial(e_d E_d)}{\partial E_d}}\right\} = \tag{22}$$

$$= \frac{\partial U_d}{\partial I_d}\left\{1 + \frac{1}{\mu_d} + \frac{\rho}{\mu_d}\frac{\partial\mu_d}{\partial\rho}\right\}\frac{\partial(e_d E_d)}{\partial E_d}. \tag{23}$$

---

[6] The expression for $\frac{\partial}{\partial\rho}\frac{1}{k_p}\frac{\mathrm{d}U_d}{\mathrm{d}E_d}$ actually independent of this specification. It, however, equals $\frac{1}{k_p}\frac{\partial}{\partial\rho}\frac{\mathrm{d}U_d}{\mathrm{d}E_d}$ only if the above choice of the variable parameter is made; and it is the sign of $\frac{1}{k_p}\frac{\partial}{\partial\rho}\frac{\mathrm{d}U_d}{\mathrm{d}E_d}$ that is crucial for the sign of $\frac{\mathrm{d}E_d^*}{\mathrm{d}\rho}$.

Finally, noting that $\eta_d = \frac{\partial e_d}{\partial E_d}\frac{E_d}{e_d}$, one obtains

$$\frac{\partial}{\partial \rho}\frac{1}{k_p}\frac{\mathrm{d}U_d}{\mathrm{d}E_d} = \frac{\partial U_d}{\partial I_d}\left\{1 + \frac{1}{\mu_d} + \frac{\rho}{\mu_d}\frac{\partial \mu_d}{\partial \rho}\right\}(\eta_d + 1)e_d, \tag{24}$$

which is equation (16) in the main text.

Appendix B

$$\rho_{\min}(1+\mu_d)_{\rho=\rho_{\min}} = \frac{1}{k_p[(\eta_d+1)e_d]_{E_d=0}} = \rho_{\max}(1+\mu_d)_{\rho=\rho_{\max}} \tag{25}$$

implying

$$\frac{(1+\mu_d)_{\rho=\rho_{\min}}}{(1+\mu_d)_{\rho=\rho_{\max}}} = \frac{\rho_{\max}}{\rho_{\min}} \tag{26}$$

Additionally there must occur a sign change of equation (16),

$$\frac{\rho}{\mu_d}\frac{\partial \mu_d}{\partial \rho} < -(1+\mu_d) \tag{27}$$

For $\rho > \rho 0$, necessitating a sufficiently fast decrease of μ_d with increasing ρ. In any case, it must be pointed out that ρ cannot become arbitrarily large. L_d must stay positive implying[7]

$$(T - E_d)/k_p - (S_p + \rho S_d) > 0 \tag{28}$$

Already in the absence of a solution of the first-order condition, i.e., while E_d^* is still zero, the work performed in the two jobs, $S_p + \rho S_d$, can violate this condition. In that case there would be no selection effort at any value of ρ.

[7] As $\mu_d \to 0$ for increasing ρ, any additional activity is very costly in terms of effort for the DP and therefore will not take place.